

Deliverable 2.1 Initial project results summary for Exploitation

Maj Institute of Pharmacology, Polish Academy of Sciences

The transcript variation analysis set was prepared as follows:

Brain slices containing the striatum and prefrontal cortex (Bregma 1.34-0.74, based on Paxinos & Franklin, 2001), as well as the thalamus (Bregma -0.94 - -1.82, based on Paxinos & Franklin, 2001) were cut on a vibratome VT1200 (Leica) into 200µm sections. Additionally, the thalami were dissected from the slices under stereo microscope (StemiDV4, Carl Zeiss, Germany) using sterile needles. RNA was isolated using the single-step acid guanidinium thiocyanate-phenolchloroform extraction method (Chomczynski, 1993).

Reverse transcription was performed using the Omniscript RT Kit (Qiagen Inc., USA). Sequencing was performed by Novogene UK using the Nanopore PromethION platform. Each sample was sequenced three times. The number of sample bases varied between 5.2 and 11.5 Gbp per sample, while the number of reads fluctuated between 3.8 and 8.2×10^6 . The median read length in all samples was relatively consistent and oscillated between 1 002 and 1 165 bp.

The length of the shortest read in the subset of the longest sequences that together represent $\geq 50\%$ of the nucleotides in the sample varied from 1 875 to 2 172 bp with higher values obtained in the samples containing the striatum and prefrontal cortex than those from the thalamus. The median read quality was high and consistent between samples, reaching values between 12.1-13.0.

The percentage of reads with relatively low quality (<7) did not exceed 3.5% in any of the samples. All FASTQ files for each brain region were merged and the data were aligned to the GRCm38_102/mm10 genome with the Minimap2 program (v. 2.24). Alignment results were filtered for strand mismatches and invalid 3' ends. The final number of valid alignments was ~ 35 million.

Data is uploaded to the database of the National Library of Medicine ([National Center for Biotechnology Information \(nih.gov\)](https://www.ncbi.nlm.nih.gov/bioproject/1032908)) and could be found under the link:

[ID 1032908 - BioProject - NCBI \(nih.gov\)](https://www.ncbi.nlm.nih.gov/bioproject/1032908) – general information

Direct links:

1. [Long-read nanopore analysis of Mus musculus \(C57BL/6\): male coronal sections \(striatum\)](#)
2. [Long-read nanopore analysis of Mus musculus \(C57BL/6\): female coronal sections \(striatum\)](#)
3. [Long-read nanopore analysis of Mus musculus \(C57BL/6\): male coronal sections \(thalamus\)](#)
4. [Long-read nanopore analysis of Mus musculus \(C57BL/6\): female coronal sections \(thalamus\)](#)